Supplementary Information

Physics-guided surrogate learning enables zero-shot control of turbulent wings

Yuning Wang[1*], Pol Suárez[2], Mathis Bode[3], Ricardo Vinuesa[1,2*]

[1*]Department of Aerospace Engineering, University of Michigan, Ann Arbor, 48109, Michigan, USA.
[2]FLOW, Engineering Mechanics, KTH Royal Institute of Technology, Stockholm, SE-100 44, Sweden.
[3]Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, Jülich, 52425, Germany.

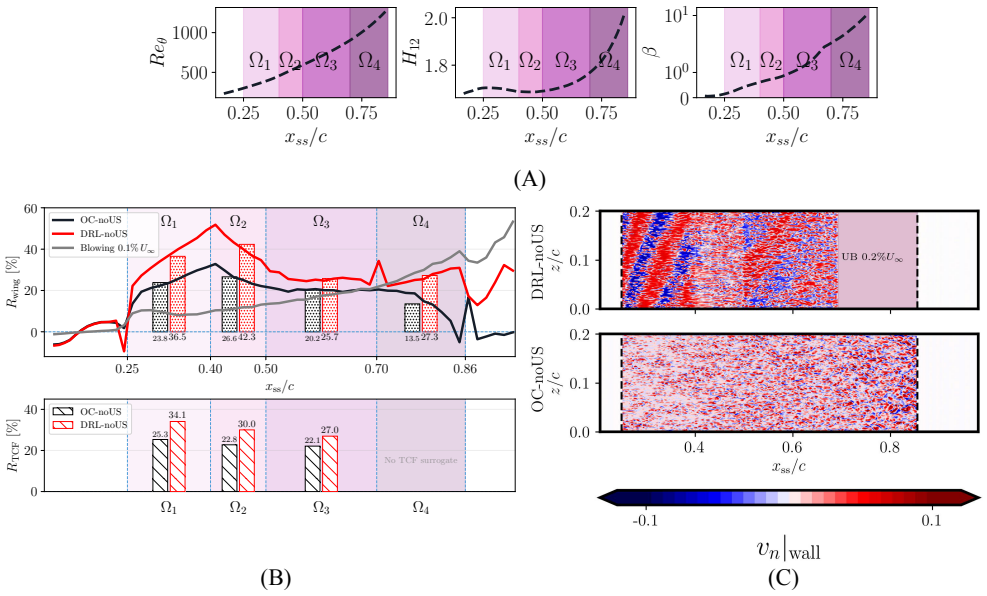
*Corresponding author(s). E-mail(s): yuninw@umich.edu; rvinuesa@umich.edu;
Contributing authors: polsm@kth.se; m.bode@fz-juelich.de;

Supplementary information 1: Control impact on aerodynamic efficiency

In the main text, the control impact on the total drag coefficient budget was reported. Table 1 presents the lift and drag coefficients for all cases together with their relative changes with respect to the uncontrolled base flow. The lift coefficient is defined as the lift force per unit span, $C_l = f_l/(bq)$, and the lift-to-drag ratio is accordingly given by $L/D = C_l/C_d$.

The reported aerodynamic coefficients reported are evaluated using a backward-cumulative averaging procedure following [1]. Specifically, for each case the time-averaged statistics are computed over a window starting at T_{FO}, which is selected as the time that minimises the run-to-run variation of the backward-cumulative aerodynamic efficiency L/D. This criterion ensures that the initial transient response to control activation is excluded from the average, while retaining as long a statistically stationary record as possible. A uniform value of $T_{\mathrm{FO}} = 6$ flow-over times is adopted across all cases, which falls within the low-variation window identified by the convergence analysis for each individual simulation. We note that the strict per-case optimum of T_{FO} varies slightly (between $t = 2$ and $t = 5$ depending on the case and the quantity considered); the choice of a common $T_{\mathrm{FO}} = 6$ is therefore conservative in the sense that it excludes slightly more of the early development than strictly necessary, and the resulting averages are not sensitive to this choice within the stationary window.

Both the DRL and OC configurations yield increases in the lift coefficient in addition to drag reduction. However, enhanced lift is not unconditionally beneficial from an aerodynamic performance perspective: for a finite wing operating at a fixed angle of attack, increased lift is accompanied by a corresponding rise in induced drag. Drag reduction, by contrast, constitutes an unambiguous aerodynamic benefit regardless of the underlying assumptions, which motivates the emphasis placed on drag modification throughout the main text.



Supplementary Figure 1 Zero-shot control performance on the BASE flow configuration. (A) Spatial profiles of Re_θ, H_{12} and β as a function of chordwise coordinate x_{ss}/c for the BASE case. (B) Relative changes in drag coefficients. (C) Instantaneous wall-normal actuation $v_n|_{wall}$ for OC and DRL at a representative time, with the controlled region indicated by dashed lines. Red and blue denote blowing and suction, respectively.

Supplementary information 2: Zero-shot deployment on the uncontrolled base flow

The method was initially applied to the base-flow configuration, in which the block-averaged $\langle Re_\theta \rangle_x$ was matched by surrogate channel flows. The profiles of Re_θ, H_{12} and β of the base flow are shown in Supplementary Figure 1(A). The rationale for choosing Re_θ as the matching target is that it characterizes TBL development under mild adverse pressure gradients; this approach also achieved promising performance on the suction side of a NACA0012 at the same Re_c and $\alpha = 0°$, as reported in [2]. It is also worth noting that matching H_{12} alone is insufficient for Ω_4, where the very strong APG and attendant flow history limit the improvement that any surrogate-based strategy can achieve.

Supplementary Figure 1(B) shows the achieved drag reduction R:

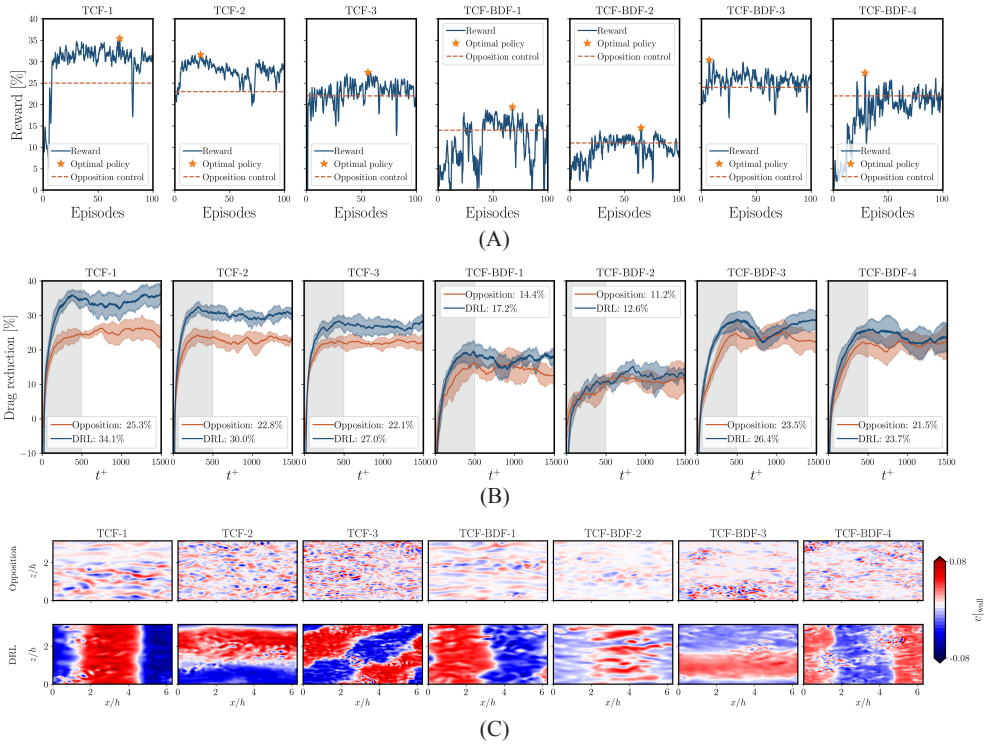
$$R = 1 - c_f^{\mathrm{ctrl}}/c_f^{\mathrm{ref}}, \tag{1}$$

where the superscripts "ctrl" and "ref" denote the controlled and its uncontrolled states, respectively.

Consistent with the main results, DRL consistently outperforms OC in both the channel and wing environments. The motivation for imposing uniform blowing at

$0.2\%\,U_\infty$ in the no-suction configuration is to achieve drag reduction comparable to that obtained with uniform blowing at $0.1\%\,U_\infty$ applied throughout the control area, as clearly supported by the wing results. On the other hand, applying uniform blowing alone increases the C_d by drastically increasing $C_{d,p}$ (see Table 1).

Note that resorting to a non-DRL policy in Ω_4 does not diminish the novelty of the framework; rather, following the block-partitioning strategy, the approach can be extended as a collection of optimal control methods applied independently to each block. Combining OC with $0.2\%\,U_\infty$ uniform blowing does not outperform the current hybrid configuration because the drag reduction is lower in the upstream blocks. The instantaneous wall actuation is shown in Supplementary Figure 1(C); transferable patterns in Ω_1–Ω_3 are identifiable, consistent with the results in Supplementary Figure 2(C).



Supplementary Figure 2 Results of exploration and exploitation in TCF surrogates. (A) Reward evolution during training. Curves are smoothed using sliding-window averaging with a window size equal to one episode. (B) Drag-reduction rate during exploitation. The solid line and shaded band denote the mean and standard deviation across six initial conditions. The gray region indicates the initial transient ($t^+ \leq 500$); reported values are computed as temporal averages after removing this transient. (C) Instantaneous wall actuation for OC and DRL at a representative time.

Case	C_l	$C_{d,f}$	$C_{d,p}$	$C_d = C_{d,f} + C_{d,p}$	L/D
Base Flow	0.867	0.0128	0.0087	0.0215	41.0
OC-noUS	0.873 (+1.0%)	0.0121 (-4.7%)	0.0083 (-4.6%)	0.0204 (-5.1%)	42.8 (+4.4%)
DRL-noUS	0.858 (-1.0%)	0.0118 (-7.8%)	0.0085 (-2.3%)	0.0203 (-5.6%)	42.3 (+3.1%)
BASE-UB 0.1%U_∞	0.833 (-3.9%)	0.0122 (-4.7%)	0.0099 (+13.8%)	0.0221 (+2.8%)	37.7 (-8.1%)
BASE-US 0.2%U_∞	0.925 (+6.7%)	0.0140 (+9.4%)	0.0066 (-24.1%)	0.0206 (-4.2%)	44.9 (+9.5%)
OC-US 0.2% U_∞	0.944 (+8.8%)	0.0134 (+4.7%)	0.0059 (-32.2%)	0.0193 (-10.2%)	48.9 (+19.3%)
DRL-US 0.2% U_∞	0.942 (+8.6%)	0.0130 (+1.6%)	0.0062 (-28.7%)	0.0192 (-10.7%)	49.1 (+19.6%)

Supplementary Table 1 Control impact on aerodynamic efficiency. Integrated lift (C_l), pressure-drag contribution ($C_{d,p}$), skin-friction-drag contribution ($C_{d,f}$), total drag (C_d), and aerodynamic efficiency (L/D) for all cases considered. Values in parentheses denote the relative change with respect to the uncontrolled reference case.

Supplementary information 3: Exploration and exploitation in channel-flow environments

Independent MARL policies are trained for each TCF surrogate corresponding to the wing configuration. Supplementary Figure 2(A) shows the reward evolution during training for all surrogates. The reward does not evolve monotonically across episodes, reflecting a known characteristic of TD3 in which policy updates can temporarily degrade performance before recovering. This makes early stopping a non-trivial decision, as it is generally unclear whether the best policy has already been encountered or whether further exploration would yield improvements. The optimal policy is therefore selected as the one achieving the highest reward over the full training history, with convergence typically reached within 70 episodes across all surrogates.

Supplementary Figure 2(B) reports the instantaneous drag-reduction rate as a function of t^+ during exploitation with the control policy fixed, averaged across six independent TCF realizations and excluding an initial transient of $t^+ < 500$ [3, 4]. The DRL policies achieve higher drag reductions, consistently outperforming OC for the same blocks. OC results are in good agreement with prior studies at comparable Re_τ [5, 6], providing confidence in the fidelity of the surrogates and implementation of our framework.

Examining the spatial distribution of wall actuation provides additional insight into how individual agents cooperate to achieve a global drag-reduction objective, information not accessible from conventional turbulence statistics owing to flow homogeneity. The DRL policies learn actuation patterns that differ markedly from OC, which prescribes wall blowing and suction strictly proportional to the wall-normal velocity fluctuations at the sensing plane.

References

[1] Atzori, M. *et al.* Aerodynamic effects of uniform blowing and suction on a NACA4412 airfoil. *Flow Turbul. Combust.* **105**, 735–759 (2020).

[2] Lagemann, C. *et al.* HydroGym: a reinforcement learning platform for fluid dynamics (2025). Preprint at https://arxiv.org/abs/2512.17534, arXiv:2512.17534.

[3] Guastoni, L., Rabault, J., Schlatter, P., Azizpour, H. & Vinuesa, R. Deep reinforcement learning for turbulent drag reduction in channel flows. *Eur. Phys. J. E* **46**, 27 (2023).

[4] Beneitez, M., Cremades, A., Guastoni, L. & Vinuesa, R. Improving turbulence control through explainable deep learning (2025). Preprint at https://arxiv.org/abs/2504.02354, arXiv:2504.02354.

[5] Stroh, A., Frohnapfel, B., Schlatter, P. & Hasegawa, Y. A comparison of opposition control in turbulent boundary layer and turbulent channel flow. *Phys. Fluids* **27**, 075101 (2015).

[6] Yao, J., García, E. & Hussain, F. Drag reduction via opposition control in turbulent channel flows at high Reynolds numbers. *Phys. Rev. Fluids* **10**, 094604 (2025).